Nonsilicon, Non-von Neumann Computing—Part I

By SANKAR BASU, *Fellow IEEE*
Guest Editor

RANDAL E. BRYANT, *Fellow IEEE*
Guest Editor

GIOVANNI DE MICHELI, *Fellow IEEE*
Guest Editor

THOMAS THEIS, *Fellow IEEE*
Guest Editor

LLOYD WHITMAN
Guest Editor

I. BACKGROUND

The future of computing is at crossroads. The technological advances that have sustained the exponential growth of computing performance over the last several decades are slowing and the roadmap for future advances is uncertain. The phenomenal expansion of computing power has made computers ubiquitous, spawned a $300 billion semiconductor industry, enabled unprecedented global economic growth, and transformed many aspects of society at large. Emerging technologies are placing an ever-growing and changing demand on computing, especially the profusion of data from the Internet of Things, large-scale scientific experiments (high-energy physics, astronomy, and genomics), autonomous vehicles, social media (including video), national security systems, and the finance sector. Transmitting, storing, processing, and analyzing this data explosion with the requisite speed and performance—and enabling significant processing and analysis to occur locally or at network nodes (i.e., edge computing)—may require a radical departure from the traditional computing paradigm of von Neumann computing architectures running on CMOS-based digital logic. New paradigms will likely require a range of new devices, software, design and simulation tools, and benchmarking, and may ultimately require rethinking the tasks that computing machines undertake. Recently, government, industry, and academia collectively have recognized that the future of computing requires a new, multidisciplinary research and development agenda.

This month's special issue provides insight into recent technological developments in the exciting field of "beyond Moore computing."

The technology challenges facing the semiconductor industry have widely been discussed [1]. The industry itself has described the most pressing needs in a September 2015 report, *Rebooting the IT Revolution: a Call for Action* [2], and proposed a new research agenda and path forward in *Semiconductor Research Opportunities: An Industry Vision and Guide* [3]. The collective message of these reports is the need for a body of fundamental research to create a new, more energy-efficient computational engine enabling real-time analysis of extraordinary volumes of high-rate data streams generated by the emerging information technology applications. While the technical issues surrounding the problem have been discussed in the research community for some time [4], both industry and government have recently undertaken significant activities supporting cross disciplinary research to address the future of computing.

The U.S. National Nanotechnology Initiative (NNI) in 2015 laid out a grand challenge to the nanotechnology, high-performance

Digital Object Identifier 10.1109/JPROC.2018.2884780

computing, and neuroscience communities to work together to create a new type of computer that can proactively interpret and learn from data, solve unfamiliar problems using what it has learned, and operate with the energy efficiency of the human brain [5]. U.S. Federal agencies supporting related research outlined the emerging and innovative solutions needed to realize this vision of future computing in a publicly available white paper [6], and numerous meetings and workshops have subsequently focused on research addressing this challenge [7].

Multiple funding agencies in the United States have been partnering with the semiconductor industry to support innovations in computing, including the Energy Efficient Computing from Devices to Architectures (E2CDA) program at the National Science Foundation [8], the nanoelectronic Computing REsearch (nCORE) program at the National Institute of Standards and Technology [9], the Joint University Microelectronics Program (JUMP) at the Defense Advanced Research Projects Agency [10], and the Machine Intelligence from Cortical Networks program at the Intelligence Advanced Research Projects Activity [11].

The broader research community, as represented by professional societies, has also been focusing on this problem, with one of the most developed efforts being the IEEE Rebooting Computing Initiative [12]. It is our understanding that similar efforts are underway outside the United States as well—in academia, industry, and European funding agencies (e.g., European Research Council, Deutsche Forschungsgeminschaft, etc.). These efforts are also generating significant amounts of novel research on new forms of computing.

The European Union (EU) launched the Graphene Flagship initiative (coordinated by Chalmers University in Sweden) in 2013 to study nonsilicon layered materials (2-D) exemplified by graphene, MoS_2, and WSe_2 with the objectives of assessing their potentials and developing reliable and reproducible means to manufacture devices on a large industrial scale. The ultimate goal is to integrate 2-D devices and components into systems capable of providing new functionalities and application areas [13]. In 2018, the EU launched the Flagship on Quantum Technologies (coordinated by Delft University in The Netherlands), researching materials, devices, and systems for quantum computing [14]. Interestingly enough, also the EU Flagship on Human Brain in Neuroscience (coordinated by EPFL Switzerland) has a significant component devoted to neuromorphic computation [15]. ERC Flagship initiatives call for the collaboration of many partners with a total budget of €1 billion each and a ten-year lifespan.

II. TECHNICAL SCOPE OF THE ISSUE

The advances in computing associated with Moore's Law and Dennard scaling—the drive toward ever smaller and faster transistors—have slowed in recent years as some fundamental limits of the long-established silicon CMOS technology have been approached. In particular, clock frequencies have been stagnant since roughly 2005, limited by the unavoidable tradeoff between power dissipation and switching speed. Thus, in a search of more power-efficient computing, a plethora of strategies and techniques for low-power computing involving all layers of the computing stack—from materials to devices, circuits, architectures, software, and application programs—are being considered.

New materials and substrates for building faster switching devices, possibly with lower power dissipation, have been investigated for the past several years. Several novel materials, including carbon nanotubes and graphene, and other 2-D materials such as MoS_2 have been identified as promising candidates for this purpose. On the other hand, switching devices which represent digital state by some physical property other than charge or voltage are also of interest. These devices exploit "alternative state vectors" such as magnetic field, electron spin, optical excitations, and phase state in various materials. Attempts to explore new materials are not only confined to the search for faster, lower power switching devices but also include a need for improved interconnects, which are also a significant contributor to the power dissipation problem.

While new device concepts offer the potential for significant advances, novel computing architectures to augment or supplant the traditional von Neumann architecture constitute a complimentary approach. Such approaches include processor-in-memory architectures, which alleviate the computational bottleneck associated with transferring data from memory to computing (logic) units, along with more ambitious and futuristic architectures inspired by the human brain or drawn from statistical mechanics. The brain operates at about 20 W and yet achieves computing tasks unmatched by current computing paradigms, especially in the areas of perception and learning. While neuromorphic paradigms based on our evolving knowledge of (computational) neurosciences appear promising among the non-von Neumann approaches and much current activity is devoted toward this goal, yet other alternatives for attacking problems known to be computationally hard or intractable in theoretical computer science are also being explored by building hardware that emulates models of physics. The latter types of approaches include computational structures based on the Ising model from statistical physics or those consisting of arrays of weakly coupled nonlinear dynamical systems (e.g., oscillators), with potential for implementation in a variety of different hardware technologies.

Efforts in recent decades by material scientists and device engineers have focused on inventing better switches (i.e., transistors) without much consideration of the computer architecture that they

would ultimately support. Likewise, the research on computer architecture and software has evolved in a device agnostic fashion, which has not made use of potentials for new technologies at the device and circuit levels of the computing stack. While this development was appropriate for its time and was necessary for rapid progress of technology in the past, a close interaction between scientists and engineers across the entire spectrum of expertise resulting in a multidisciplinary effort is now mandated at the current state of the art. Thus, for example, the research in device engineering at the micrometer and nanometer scales should be integrated with concomitant work by computing architects and software engineers attempting to leverage the characteristics of those devices, with the entire effort aimed at creating a new high-performance system. Much research to fill this large, integration gap, broadly dubbed here as "beyond Moore computing," is currently being undertaken by various groups around the world. The current special issue for the PROCEEDINGS OF THE IEEE covers relevant work by the leading experts in the field broadly described above.

Of course, the scope of a program of research as broad as this may as well include several other relevant topics not explicitly mentioned here. Unsurprisingly, since the special issue was organized more than a year ago, remarkable new developments have taken place. Accordingly, Part II of this special issue has been planned and will appear approximately a year from the publication of this Part I. Continuing with the general topic of nonsilicon and non-von Neumann computing, Part II will address topics that received incomplete or no coverage at all in Part I.

Meanwhile, the hardware community has also been stirred up by the push for AI technologies as well as quantum computing and quantum information sciences. While the fields of AI and machine learning have gone through the cycles of interest in the past, the emergence of effective (and, therefore, economically important) algorithms for pattern matching and correlation finding has brought interest to an intense new peak. The need to implement deep learning algorithms with energy-efficient hardware has in turn spiked interest in neuromorphic hardware. The NNI grand challenge of 2015 presciently focused on low-power cognitive computing as needed for local or edge implementations of AI. While several papers in this special issue do indeed touch upon the topics of AI and brain-inspired computing, Part II will have more on this topic. Furthermore, two papers originally considered for publication in Part II of this special issue have recently appeared in the PROCEEDINGS OF THE IEEE as standalone publications (cf. [16] and [17]).

Inclusion of the topic of quantum computing was debated among the Guest Editors and also the larger editorial board when this special issue was originally conceived but was set aside for Part II of this special issue, which will have a separate subsection on quantum computing. Instead, the present Part I only includes a brief article "An outlook for quantum computing," by Maslov *et al.* in the Point of View (PoV) category in the PROCEEDINGS OF THE IEEE.

We group the papers appearing in Part I of this special issue into several broad categories. More detailed perspectives of Guest Editors on specific papers in this issue follow. Note that almost all of these papers survey ongoing large projects that are still rapidly evolving. Thus, the papers may not do full justice to the extent and coverage of such projects. They are simply meant to whet the appetite of the readers for further reading. Most projects also have open and easily accessible extensive websites of their own available to the readership of the journal for further reference. These are cited whenever possible.

III. NOVEL MATERIALS AND DEVICES

Our first group of papers focuses on emerging electronic materials and devices. The first paper in this category describes imperfection-immune design of carbon nanotube transistors in 3-D structures, and the second paper reviews research on a potentially very important emerging transistor concept, the negative capacitance field-effect transistor (NCFET). The third paper is on DNA storage, and although not aligned with present day mainstream computing research *per se*, it is viewed as an important and emergent theme in the computing community.

The first paper of the issue titled "The N3XT approach to energy-efficient abundant-data computing" from a group of researchers primarily from Stanford University undertakes the problem of energy-efficient computing by attacking it from a multitude of different directions. Based on the imperfection immune design of carbon nanotube transistors (CNT) developed by the group, the paper reports a 3-D heterogenous architecture particularly suitable for BIG DATA (abundant data to use their terminology) applications. Thus, it is a unique combination of emerging nonsilicon technologies (CNT in this particular case) with monolithic 3-D architectures, and tacitly makes use of other architectural techniques, e.g., near memory computation. A hundredfold performance gain is expected for various applications. While early versions of this work appeared elsewhere (e.g., patronized by the IEEE Rebooting Computing initiative and in *Nature*), the present version provides current updates at the time of its writing. We are aware that the technology is being further developed for specific applications. More details are available in [18].

In the second paper titled "Negative capacitance transistors," Wong and Salahuddin from the University of California at Berkeley provide an overview of theoretical and experimental work on this promising new type of field effect transistor. Because the NCFET integrates a thin-film ferroelectric insulator into the gate stack of an otherwise conventional FET, the switching mechanism is altered

so as to relax existing constraints on the power and performance of digital logic. At the time of the original 2008 theoretical proposal, building the proposed device with known ferroelectric materials at nanometer-scale dimensions appeared impractical. However, the subsequent discovery of new materials exhibiting strong ferroelectric behavior in films as thin as a few nanometers has sparked intense research and development efforts. It is still too early to predict when the NCFET will be commercialized and how important it will become for information technology, but recent experimental results highlighted by Wong and Salahuddin are very encouraging. Further advances in the understanding of thin-film ferroelectrics, including the contribution of polarization domain formation and domain wall motion to switching dynamics, will be essential in bringing the technology to market.

The paper titled "DNA data storage and hybrid molecular–electronic computing" was written by computer architecture researchers from the University of Washington and Microsoft Research. While DNA computing is a broader field of activity, not fully addressed in this special issue, the paper attempts to address the problem of long-term storage and retrieval of large volumes of data. Indeed, based on back of the envelop calculations, it has been claimed that a pound of DNA can, in principle, hold enough data to encapsulate all human knowledge. While this sounds impractical for various reasons from a computing perspective, including the speed of storing, reading and retrieving such data—not to speak of long-term durability depending on temperature variations and other environmental effects, computer architects are beginning to experiment with DNA technology for its potential. It appears that the synthesis and decoding of DNA molecules pose challenges very different from those of electronic systems. The authors thus argue for a hybrid molecular–electronic architecture that plays to the strengths of both. The paper first presents design of a DNA-based archival storage system, which includes the largest demonstration to date of DNA digital data storage of over three billion nucleotides encoding over 400 MB of data, and then goes on to propose a more ambitious hybrid–electronic design that uses a molecular form of near-data processing for massive parallelism presenting a model that demonstrates the potential feasibility of such systems in the near future. For further publications by the group, the reader may be referred to [19].

IV. PHYSICS-BASED NON-VON NEUMANN PARADIGM

The second group of papers discusses points of departures from conventional computing paradigms. The first of the two papers in this group considers arrays of weakly coupled oscillators for simulating dynamical systems. The second paper considers a principled approach to the design of non-von Neumann architectures by emphasizing the use of information-based metrics, and thus enabling determination of fundamental limits on energy, latency, and accuracy, and thereby guiding the exploration of statistical design principles for low signal-to-noise ratio circuits. Perhaps one can characterize the above as physics-based architectures as opposed to neurally inspired architectures that follow in the next section.

The paper titled "Computing with networks of oscillatory dynamical systems" from a group of researchers primarily from Notre Dame University and Georgia Tech delves deeper into issues of unconventional computing paradigms, via a radically different approach. Stepping back from conventional digital computers, it adopts the broader view that can be paraphrased by saying "physics computes." The broader agenda pursued at the EXCEL Center [20] at Notre Dame—an E2CDA-NRI supported group of researchers—may be thought of as simulating complex dynamical systems in electronic circuits (digital or analog), the time evolution of the dynamics of which solves the computing problem at hand. A plethora of computing problems including various pattern recognition problems of interest, queuing problems, as well as intractable combinatorial optimization problems of discrete mathematics studied in theoretical computer science can be approached in this fashion. The paper demonstrates a few specific examples by building an array of weakly coupled oscillatory networks as examples of relevant dynamical systems of interest. Coupling is achieved by appealing to novel properties of the electronic materials, e.g., transition metal oxides (vanadium oxides) in this particular case—thus demonstrating that while the architecture of the intended "computing machine" is radically non-von Neumann, material properties at the lowest level of the computing stack can be profitably leveraged if one is willing to accommodate a novel paradigm. While the paper presents the results on solution of graph coloring problems as an example of the so-called NP hard problems, and image detection as an example of a more practical application, a variety of other scenarios are also discussed elaborating on the potential of this general methodology, including their use in (stochastic) spiking neural network design. The methodology applies in analog computing scenario as well, and although not explicitly reported in this paper, other prestigious publications (e.g., *Nature*) authored by the members of the group have recently discussed alternate dimensions of this emerging paradigm.

The paper "Shannon-inspired statistical computing for the nanoscale era" is contributed by a research group from the University of Illinois at Urbana–Champaign. This paper first questions the limitations of traditional von Neumann computing due to the data-centric requirements of emerging applications as well as variability and possible failure of the underlying fabrication technology.

Indeed, the data storage and transfer requirements of effective data-centric hardware support for realizing machine learning algorithms poses severe constraints on the processor-memory interface, due to the large amount of data and its variation. Similarly, edge computing often requires acquiring a plurality of varying data from sensors, thus posing tight constraints on the sensor to processor/memory interface. The contribution focuses on the concept of statistical computing that addresses the variation in workload and in computing substrate characteristics. The statistical computing paradigm leverages the use of information-based metrics to determine bounds on energy and latency: it finds its roots in seminal work by Claude Shannon and provides a framework for composing computational and storage systems in the presence of noise. We can expect the impact of this work to grow as our computing systems will tackle increasingly larger data sets and will increasingly realize more downscaled nanodevices.

V. NEUROMORPHIC PARADIGM

Albeit an example of non-von Neumann architecture as well, the study of neuromorphic architectures has a life of its own. The next group of papers deals with this latter type of non-von Neumann architectures. The first of the three papers in this group is from an IBM Watson research group that explores the current state of neuromorphic deep learning architectures in silicon CMOS technology. The second one takes a more unconventional approach to learning machines based on little explored but much promising notions of hyper dimensional computing. The third paper provides a broad brush approach to various issues of an ongoing project on different aspects of neuromorphic computing.

The paper "The next generation of deep learning hardware: Analog computing" is a timely review of hardware implementation of the much discussed deep learning algorithm by a leading group from IBM Research Division that had much to do with the IBM Watson engine catapulting the AI technology to the forefront in recent years. This review article starts with a leisurely introduction to deep learning for the uninitiated, and argues that GPU-like, but customized hardware architectures would be needed for deep learning-based AI to reach its full potential. The experience of the research team at IBM is elaborated by bringing in in-memory compute paradigm, and the possibilities of integrating analog circuitry into deep learning networks. Roles of different types of emerging memory technologies in this scenario, e.g., phase change memory, resistive RAMs, and electrochemical devices, are discussed. While this is an evolving field of research and much is to be gleaned from current activities in various industrial labs, the paper provides an early glimpse of current buzz on customized hardware for AI applications.

The paper "Efficient biosignal processing using hyperdimensional computing: Network templates for combined learning and classification of ExG signals" is part of yet another class of hardware processing very directly inspired by the computing grand challenge mentioned at the beginning of this article. Indeed, one of the goals of this broader agenda of computing research has been to build machines that learn from very little data (such learning problems have been dubbed "one-shot learning" in the computer vision community, and are relevant in other problem domains as well). This paper is a collaborative effort between cognitive neuroscientists from the Redwood Center for Theoretical Neuroscience, and hardware engineers from Berkeley, as well as ETH Zurich. A web search with keywords "hyper-dimensional computing" would produce a plethora of information from the latter mentioned neuroscience institute for the more interested reader on this methodology. While several other applications, involving both 1-D signals and multidimensional data sets as for images, are the ultimate target of this line of research, this paper deals with EMG/EEG data for biomedical applications. More exotic problems, e.g., human speech/language recognition and visual scene understanding remains a longer term goal of this joint hardware–software–algorithmic effort. Completely novel data representations and encodings in hyper-dimensional spaces are sought here. Work is underway that may ascertain the feasibility of low-level algorithmic processing implemented with specific material substrates to have stronger promise for learning machines from little data.

The paper "Braindrop: A mixed-signal neuromorphic architecture with a dynamical systems-based programming model" has its genesis in the retina chip originally developed in Carver Mead's VLSI group at CalTech. Following up that line of work, this collaborative paper by analog and digital hardware engineers and brain scientists reports on the construction of a programmable computational machine inspired by the human brain. Braindrop is a single-core instantiation of a previous multicore version called the brainstorm chip that adopted the same philosophy from a cognitive standpoint. There exists significant difference of this approach with the regular artificial neural networks pursued in the larger community in many respects (e.g., analog/mixed signal, spiking neurons, and sparse spatial encoding) making the model closer to true biological neural networks. The class of problems that such a chip can handle may be described by multidimensional nonlinear differential equations, and the computations are described by the user in a manner agnostic to the underlying hardware. This paper described a snippet of a project that has been ongoing for several years at Stanford University [21]. While the idea provides a broader framework, several recent developments, including some industrial neuromorphic chip designs (e.g., IBM TrueNorth), have been influenced by this work. Nevertheless, some comparisons with

other spiking architectures have also been provided.

VI. CMOS AND HIGH-PERFORMANCE COMPUTING

Next, the discussion returns more or less to silicon CMOS technologies in the fourth and last section of this special issue, in which three papers are presented. Here, one of the papers is on logic synthesis with emerging technologies, and the second paper deals with an important issue of specialization in computing which can potentially provide at least a shorter term strategy to combat Moore's law slowdown. The issue ends with the vision of a supercomputer or a server in a box that elaborates on a European project that has been ongoing for some time.

The paper titled "Logic synthesis for established and emerging computing" is from a group at EPFL, Switzerland, led by one of our guest editors who is a recognized world expert in the field of logic design and provides a state-of-the-art exposition on the status of logic design flows in conventional silicon CMOS as well as several of the emerging technologies, e.g., QCAs, optical architectures, leading up to majority gates, and quantum logic gates. The paper begins with a succinct discussion of how logic synthesis, in some sense, has been a holy grail of design of computers and how a plurality of post-CMOS nanoemerging technologies continues to enforce a reorientation of classical logic synthesis methods, via altered set of various design constraints, requirements of functionality, and sometimes ill-defined objective criterion for optimization. On the one hand, the reader should be convinced that the tools and techniques of design automation, although well established in silicon CMOS technology by now, when adequately modified, can also be profitably leveraged in post-CMOS era technologies. On the other hand, the paper presents only a tip of the iceberg, and claims that much exciting future work remains to be carried out in the context of emerging technologies. In order to whet the fancy of the perceptive reader, and to place the paper in a proper perspective, it may be mentioned that by viewing logic synthesis as a complex optimization problem, currently emerging AI techniques (e.g., Google DeepMind used for AlphaGo) can be viewed as potential candidates for advancing powerful design automation tools for making them available to nonexperts in design.

The paper titled "Customizable computing—From single chip to datacenters" provides the vision of a group of researchers at the University of California at Los Angeles (UCLA). System design has always been searching for a balance between hardware and software. The RISC initiative and success of the 1980s were driven by the idea of streamlining hardware and moving functionality to software. The broad applications of field-programmable hardware are based on hardware modularity and specialization by software means. The paper argues that 10 to 100 times improvements in performance/energy figures (over general purpose processors and within specific application domains) can be achieved by the use of custom-designed accelerators. Such accelerators depart from the classical von Neumann architecture and feature heterogeneity and customization in the search for energy efficiency. A prime application of customizable computing is within data centers, where composition of custom accelerators is required to deal with large heterogeneous data sets and efficient data searching. Moreover, the paper addresses high-level synthesis with architectural template generation and efficient runtime support, thus giving an overall look at the software environment for creating and operating such programmable accelerators. It is necessary to stress the importance of this domain, as datacenter technology has strong implications on everybody's daily efficacy when browsing for information, as well as on the societal costs of running datacenters that consume a significant percentage of the world's energy budget.

This issue ends with a paper that provides a somewhat different flavor of applications of basic technologies to energy-efficient servers and high-performance computing of the future. It reports on the current state of a project under development in Dresden, Germany, under the sponsorship of the Deutsche Forschungsgemeinschaft as part of their Sonderforschungsbereich program lasting more than a decade. A goal of this evolving project on highly adaptive energy-efficient computing (HAEC) [22] is to produce, in a box of approximately one liter size, a (super)computing machine via integration of a multitude of heterogonous technologies ranging from multiple 3-D chip stacks, to near memory architectures, to biplane optical, and/or wireless interconnects. Such a machine would harness the ultimate power of CMOS technology by exploiting parallelism and runtime resource management to its fullest extent by deploying a sea (millions) of core processing units. Future goal is to incorporate emerging technologies such as terarhetz processing units as well. The authors claim that the HAEC Box is a vision that can be expected to become reality around 2035. Accordingly, base technologies for components are being developed. Its small size opens new possibilities by bringing servers closer to the user, e.g., by combining them with a base station of a cellular network or a Wi-Fi hotspot, and allowing reduction of network traffic loads and energy consumption significantly and, hence, enabling the mobile edge cloud. ■

REFERENCES

[1] H. N. Khan, D. A. Hounshell, and E. R. H. Fuchs, "Science and research policy at the end of Moore's law," *Nature Electron.*, vol. 1, pp. 14–21, Jan. 2018. [Online]. Available: https://doi.org/10.1038%2Fs41928-017-0005-9

[2] [Online]. Available: https://www.src.org/newsroom/rebooting-the-it-revolution.pdf

[3] [Online]. Available: https://www.semiconductors.org/clientuploads/Research_Technology/SIA%20SRC%20Vision%20Report%203.30.17.pdf

[4] National Research Council, *The Future of Computing Performance: Game Over or Next Level?* 2011. [Online]. Available: https://doi.org/10.17226/12980

[5] [Online]. Available: https://www.nano.gov/grandchallenges

[6] [Online]. Available: http://www.nano.gov/sites/default/files/pub_resource/federal-vision-for-nanotech-inspired-future-computing-grand-challenge.pdf

[7] [Online]. Available: https://www.nano.gov/node/1523#meetingsandworkshops

[8] [Online]. Available: https://www.nsf.gov/pubs/2017/nsf17531/nsf17531.pdf

[9] [Online]. Available: https://www.src.org/program/ncore/

[10] [Online]. Available: https://www.src.org/compete/jump/

[11] [Online]. Available: https://www.iarpa.gov/index.php/research-programs/microns

[12] [Online]. Available: http://rebootingcomputing.ieee.org/

[13] [Online]. Available: https://graphene-flagship.eu/project/roadmap/Pages/Roadmap.aspx

[14] [Online]. Available: http://europa.eu/rapid/press-release_IP-18-6205_en.htm

[15] [Online]. Available: https://www.humanbrainproject.eu/en/silicon-brains/

[16] V. Sze, Y. Chen, T. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," *Proc. IEEE*, vol. 105, no. 12, pp. 2295–2329, Dec. 2017,

[17] S. Yu, "Neuro-inspired computing with emerging nonvolatile memories," *Proc. IEEE*, vol. 106, no. 2, pp. 260–285, Feb. 2018.

[18] [Online]. Available: https://nanoheat.stanford.edu/publications/2015/energy-efficientabundant-data-computing-n3xt-1000x

[19] [Online]. Available: http://misl.cs.washington.edu/publications.html

[20] [Online]. Available: https://collectivecomputing.nd.edu/

[21] [Online]. Available: https://web.stanford.edu/group/brainsinsilicon/neuromorphics.html

[22] [Online]. Available: https://tu-dresden.de/zih/forschung/projekte/haec

ABOUT THE GUEST EDITORS

Sankar Basu (Fellow, IEEE) is a Program Director with the National Science Foundation (NSF), Computer and Information Sciences and Engineering Directorate, where he is involved in the broad areas of computing including design automation for micro- and nano-electronic systems. He was with the IBM Thomas J. Watson Research Center, a Faculty Member with the Stevens Institute of Technology, and has held visiting appointments at Ruhr University, Bochum, Germany, as an Alexander von Humboldt Fellow, and the MIT Laboratory for Information and Decision Systems, Cambridge, MA. In 2012, he served as an Embassy Science Fellow at the U.S. Embassy, Berlin, Germany on detail to the U.S. State Department.

He has authored over 100 refereed publications, a dozen patents, edited several volumes, and has given keynote talks at several international conferences. At NSF, he has been involved in several interagency programs, including the Energy-Efficient Computing: from Devices to Architectures (E2CDA) in response to the OSTP Grand Challenge on computing. He served on editorial boards of about 10 journals, including the Proceedings of the IEEE, the IEEE Press and was the Editor-in-Chief of the IEEE Transactions on Circuits and Systems. He has organized and chaired major conferences, including a NATO Advanced Study Institute on Statistical Machine Learning and Applications, resulting in a coedited book *Advances in Learning Theory: Methods, Models, and Applications*.

Dr. Basu is a Fellow of the IEEE, and a Fellow and Council Member of the American Association of the Advancement of Science (AAAS).

Randal E. Bryant (Fellow, IEEE) has been a faculty member at Carnegie Mellon University since 1984 and is currently a University Professor of Computer Science. He also served on the faculty at Caltech (1981–1984) and as a visiting research fellow at Fujitsu, in Kawasaki, Japan (1989–1990). He was Assistant Director of Information Technology Research and Development for the White House Office of Science and Technology Policy (2014–15), where, together with Lloyd Whitman, he architected the OSTP Computing Grand Challenge.

Dr. Bryant is well known for his research on simulating and formally verifying hardware systems. He developed switch-level simulation, a method of modeling MOS circuits using simple switches. He also developed the ordered binary decision diagram data structure, widely used in hardware synthesis, testing, and verification. Dr. Bryant is a Fellow of the Association for Computing Machinery (ACM) and a member of the National Academy of Engineering and the American Academy of Arts and Science. He was a recipient of the 1997 ACM Kanellakis Theory and Practice Award (shared with Edmund M. Clarke, Ken McMillan, and Allen Emerson) for contributing to the development of symbolic model checking, the 1989 IEEE W.R.G. Baker Prize for the best paper appearing in any IEEE publication during the preceding year, the 2010 ACM/IEEE A. Richard Newton Technical Award in Electronic Design Automation, the 2009 EDAC/IEEE Phil Kaufman Award, and the 2007 IEEE Emmanuel R. Piore Award, reflecting the impact his work has had on tools used by the semiconductor industry for verifying their designs prior to manufacture.

Giovanni De Micheli (Fellow, IEEE) is a Professor and Director of the Institute of Electrical Engineering at EPFL Lausanne, Lausanne, Switzerland.

Prof. De Micheli is a Fellow of the Association for Computing Machinery (ACM), a member of Academia Europaea, and a Foreign Honorary member of the American Academy of Arts and Sciences. He is the recipient of the 2016 IEEE/CS Harry Goode award for seminal contributions to design and design tools of networks on chips; the 2016 EDAA Lifetime Achievement Award; the 2012 IEEE/CAS Mac Van Valkenburg award for contributions to theory, practice, and experimentation in design methods and tools; and the 2003 IEEE Emanuel Piore Award for contributions to computer-aided synthesis of digital systems. He also received the Golden Jubilee Medal for outstanding contributions to the IEEE Circuits and Systems (CAS) Society in 2000; the D. Pederson Award for the best paper in the IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD/ICAS) in 1987 and 2018; and several Best Paper Awards, including in DAC (1983 and 1993), DATE (2005), Nanoarch (2010 and 2012), and Mobihealth (2016). He has been serving IEEE in several capacities, namely, Division 1 Director (2008–2009), co-founder and President Elect of the IEEE Council on EDA (2005–2007), President of IEEE CAS (2003), and Editor-in-Chief of the IEEE Transactions on CAD/ICAS (1997–2001). He has been Chair of several conferences, including Memocode (2014), DATE (2010), pHealth (2006), VLSI SOC (2006), DAC (2000), and ICCD (1989).

Thomas Theis (Fellow, IEEE) is currently IBM Distinguished Research Staff Member Emeritus with the IBM Thomas J. Watson Research Center.

Dr. Theis joined the Watson Research Center in 1978 to study the electronic properties of materials. He held various senior management and executive positions from 1984 to 2016. In 1990, he joined Silicon Science and Technology as a Senior Manager where he coordinated the transfer of copper interconnection technology from IBM Research to the IBM Microelectronics Division. As IBM's Strategist for exploratory research world-wide from 1998 to 2012 and as Director of physical sciences from 1998 to 2010, he conceived and initiated successful research programs in silicon nanophotonics and Josephson junction-based quantum computing, and championed research in nanoelectronics, exploratory memory devices, and applications of information technology to address societal needs in energy, infrastructure, and the environment. From 2010 to 2012, he organized research projects aimed at greatly improved energy efficiency in computing. From 2012 to 2016, he led the Semiconductor Research Corporation's Nanoelectronics Research Initiative. In 2016, he joined Columbia University, New York, NY, USA, as a Professor of practice in electrical engineering and the Executive Director of the Columbia Nano Initiative. In 2018 he retired from Columbia University to focus on independent research, technical consulting, and various professional activities.

Dr. Theis is a Fellow of the American Physical Society. He was the recipient of the 2015 George E. Pake Prize of the APS.

Lloyd Whitman is currently the Chief Scientist at the National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA and the Principal Assistant Director for physical sciences and engineering at the White House Office of Science and Technology Policy (OSTP), Washington, DC, USA, where he advises on a broad range of science and technology areas and U.S. research and development programs and strategies.

Dr. Whitman received the B.S. degree in physics from Brown University, Providence, RI, USA, and the M.S. and Ph.D. degrees in physics from Cornell University, Ithaca, NY, USA. He was with the U.S. Naval Research Laboratory where he led a diverse portfolio of research studying nanostructures and their integration into advanced sensor systems. He was the founding Deputy Director of the NIST Center for Nanoscale Science and Technology. He also served as the Assistant Director for nanotechnology and advanced materials at OSTP for the last two years of the Obama administration. He has authored over 160 publications and multiple patents and has been recognized with numerous media citations and awards.